LaX: Boosting Low-Rank Training of Foundation Models via Latent Crossing



Ruijie Zhang, Ziyue Liu, Zhengyang Wang, Zheng Zhang

University of California, Santa Barbara

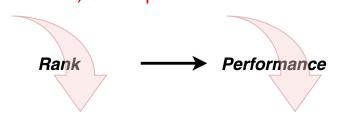


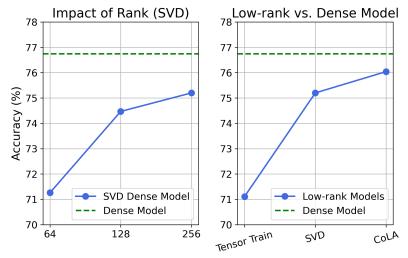




Motivation:

Low-rank models are smaller in size but: Smaller rank, worse performance.



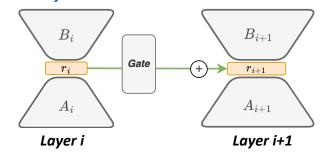


Goal:

Full-rank Performance but Low-rank models?

Our Solution: Latent Feature Fusion

Inter-layer LaX



Intra-layer LaX

Dense Layers

outputs

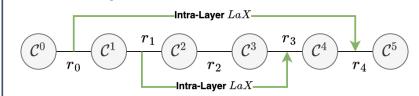
dense layer weights W.

outputs

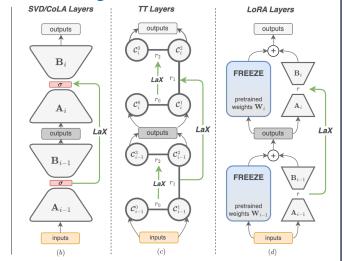
dense laver

weights $\hat{\mathbf{W}}_{i-1}$

inputs

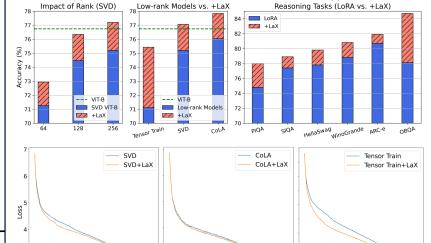


Plug to win!



Experiments:

 ViTs Pre-training (ImageNet-1k) & LoRA Finetuning (Commonsense)



LLMs Pre-training (C4)

100 150 200 250 300 0 50 100 150 200 250 300 Epoch

